International Journal of Mathematics and Computer Applications Research (IJMCAR)
ISSN(P): 2249-6955; ISSN(E): 2249-8060
Vol. 4, Issue 1, Feb 2014, 19-30

STUDY ON APPROXIMATE GRADIENT PROJECTION (AGP) PROPERTY IN NONLINEAR PROGRAMMING

ROHIT SARIN[1], B. K. SINGH[2] & SANJEEV RAJAN[3]

[1]Research Scholar, IFTM University, Moradabad, Uttar Pradesh, India

[2]HOD, Department of Mathematics, IFTM University, Moradabad, Uttar Pradesh, India

[3] HOD, Department of Mathematics, Hindu college, Moradabad, Uttar Pradesh, India

ABSTRACT

In this paper, we introduce an optimality condition that, roughly formulated, says that an approximate gradient projection tends to zero. For this reason, we call it approximate gradient projection (AGP) property. We prove that the AGP property is satisfied by local minimizers of constrained optimization problems independently of constrained qualifications and using only first – order differentiability. Unlike the KKT conditions, the new condition is satisfied by local minimizers of nonlinear programming problems, independently of constraint qualifications.

KEYWORDS: Optimality Conditions, Karush – Kuhn – Tucker Conditions, Minimization Algorithms, Constrained Optimization

INTRODUCTION

Any optimization problem essentially consists of an objective function. Depending upon the nature of objective function (O.F.), there is a need to either maximize or minimize it. If certain constraints are imposed, then it is referred to as constrained optimization problem. In the absence of any constraint, it is an unconstrained problem. Linear programming (LP) methods are useful for the situation when O.F. as well as constraints are the linear functions. Such problems can be solved using the simplex algorithm. Nonlinear programming (NLP) is referred to the followings:(a) Nonlinear O.F. and linear constraints(b) Nonlinear O.F. and nonlinear constraints(c) Unconstrained nonlinear O.F. Since the 1960s, much effort has gone into the development and application of numerical algorithms for solving problems in the two areas of optimization and systems of equations. As a result, many different ideas have been proposed for dealing efficiently with (for example) severe nonlinearities and/or very large numbers of variables. Libraries of powerful software now embody the most successful of these ideas, and one objective of this volume is to assist potential users in choosing appropriate software for the problems they need to solve. More generally, however, these collected review articles are intended to provide both researchers and practitioners with snapshots of the `state-of-the-art' with regard to algorithms for particular classes of problem. These snapshots are meant to have the virtues of immediacy through the inclusion of very recent ideas, but they also have sufficient depth of field to show how ideas have developed and how today research questions have grown out of previous solution attempts.

A Variety of applications in engineering, decision science, and operations research have been formulated as constrained nonlinear programming problems (NLPs). Such applications include neural-network learning, digital signal and image processing, structural optimization, engineering design, computer-aided-design (CAD) for VLSI, database design and processing, nuclear power plant design and operation, mechanical design, and chemical process control[1][2]. Due to the availability of a lot of unconstrained optimization algorithms, many real applications that are inherently

nonlinear and constrained have been solved in various unconstrained forms. Optimal or good solutions to these applications have significant impacts on system performance, such as low-cost implementation and maintenance, fast execution, and robust operation. Every constrained NLP has three basic components: a set of unknowns or variables to be determined, an objective function to be minimized or maximized, and a set of constraints to be satisfied. Solving such a problem amounts to finding values of variables that optimize (minimize or maximize) the objective function while satisfying all the constraints.

When the constraint qualification is the one introduced by **Mangasarian and Fromovitz [1]**, we say that the Fritz – John optimality condition is satisfied. Perhaps separation techniques are the most powerful tools for proving optimality conditions **[2 – 6]**. Therefore, according to that model global theorem, the algorithmic sequence might converge to points where the constrained qualification is not satisfied.

Therefore, both practical experience and common sense indicates that the set of feasible limit points of good algorithms for constrained optimization, although being larger than the reputation of the KKT conditions lies not only in being necessary optimality conditions under constraint qualifications, but also in the fact that they are sufficient optimality conditions if the nonlinear program is convex. Essentially, we prove that this is also true for the AGP property.

THE NECESSARY CONDITION

Consider the NLP

$$\min f(x), \text{ s.t. } x \in \Omega, \tag{1}$$

where

$$\Omega = \{x \in \mathbb{R}^n | g(x) \le 0, h(x) = 0\}, \tag{2}$$

$f: \mathbb{R}^n \to \mathbb{R}, h: \mathbb{R}^n \to \mathbb{R}^m, g: \mathbb{R}^n \to \mathbb{R}^p,$ and all the functions have continuous first partial derivatives.

Let $\gamma \in [-\infty, 0]$. For all $x \in \mathbb{R}^n$, we define $\Omega\,(x, \gamma)$ as the set of points $z \in \mathbb{R}^n$ that satisfy

$$g_i(x) + g_i'(x)(z - x) \le 0, \qquad \text{if } \gamma < g_i(x) < 0, \tag{3}$$

$$g_i'(x)(z - x) \le 0, \qquad \text{if } g_i(x) \ge 0, \tag{4}$$

$$h'(x)(z - x) = 0. \tag{5}$$

The set $\Omega\,(x, \gamma)$, a closed and convex polyhedron, can be interpreted as a linear approximation of the set of points $z \in \mathbb{R}^n$ that satisfy

$$h(z) = h(x),$$

$$g_i(z) \le g_i(x), \qquad \text{if } g_i(x) \ge 0,$$

$$g_i(z) \le 0, \qquad \text{if } g_i(x) \in (\gamma, 0).$$

Observe that

$$\Omega(x, 0) = \{z \in \mathbb{R}^n | h'(x)(z - x) = 0, g_i'(x)(z - x) \le 0, \text{if } g_i(x) \ge 0, i = 1, \dots\dots, p\}.$$

For all $x \in \mathbb{R}^n$, we define

$$d(x,\gamma) = P_{\Omega(x,\gamma)}\big(x - \nabla f(x)\big) - x, \tag{6}$$

Where $P_c(y)$ denotes the orthogonal projection of y onto C for all $y \in \mathbb{R}^n, C \subset \mathbb{R}^n$ closed, convex. The vector $d(x,\gamma)$ will be called the approximate gradient projection.

We will denote $\|\cdot\| = \|\cdot\|_2$ and

$B\ (x,\rho) = \{z \in \mathbb{R}^n |\ \| z - x \| \leq \rho\}$,

for all $x \in \mathbb{R}^n$, $\rho > 0$. As usual, for all $v \in \mathbb{R}^n$, $v = (v_1, \dots\dots, v_n)$, we denote

$v_+ = (\max\{0, v_1\}, \dots, \max\{0, v_n\})$.

Theorem 2.1

Assume that x^* is a local minimizer of (1)–(2), and let $\gamma \in [-\infty, 0]$, $\in$, $\delta > 0$ be given. Then, there exists $x \in \mathbb{R}^n$ such that

$\| x - x^* \| \leq \delta$ and $\| d(x,\gamma) \| \leq \epsilon$.

Proof

Let $\rho \in (0, \delta)$ be such that x^* is a global minimizer of *f(x)* on $\Omega \cap B\ (x^*, \rho)$. For all $x \in \mathbb{R}^n$, define

$\varphi(x) = f(x) + (\in/2\rho) \| x - x^* \|^2$.

Clearly, x^* is the unique global solution of

$$\min \varphi(x), \quad \text{s.t. } x \in \Omega \cap B\ (x^*, \rho). \tag{7}$$

For all $x \in \mathbb{R}^n$, $\mu > 0$, define

$\Phi_\mu(x) = \varphi(x) + (\mu/2)[\| h(x) \|^2 + \| g(x)_+ \|^2]$.

The external penalty theory (see for instance Ref. 12) guarantees that, for μ sufficiently large, there exists a solution of

$$\min \Phi_\mu(x), \quad \text{s.t. } x \in B\ (x^*, \rho), \tag{8}$$

that is as close as desired to the global minimizer of $\varphi(x)$ on $\Omega \cap B\ (x^*, \rho)$. So, for μ large enough, there exists a solution x_μ of (8) in the interior of $B\ (x^*, \rho)$. Therefore,

$\nabla \Phi_\mu(x_\mu) = 0$.

Thus, writing for simplicity $x = x_\mu$, we obtain

$0 = \nabla \Phi_\mu(x)$

$= \nabla \varphi(x) + \mu\left[h'(x)^T h(x) + \sum_{g_i(x)>0} g_i(x) \nabla g_i(x)\right]$.

So,

$$\nabla f(x) + \mu\left[h'(x)^T h(x) + \sum_{g_i(x)>0} g_i(x) \nabla g_i(x)\right] + (\in/\rho)(x - x^*) = 0. \tag{9}$$

Since $x = x_\mu$ lies in the interior of the ball, we have that

$\|x - x^*\| < \rho < \delta.$

So, by (9),

$\|\nabla f(x) + \mu[h'(x)^T h(x) + \sum_{g_{i(x)>0}} g_i(x)\nabla g_i(x)]\| \leq \epsilon.$

So,

$\|[x - \nabla f(x)] - (x + \mu[h'(x)^T h(x) + \sum_{g_{i(x)>0}} g_i(x)\nabla g_i(x)])\| \leq \epsilon.$

Taking projections onto $\Omega(x, \gamma)$, this implies that

$$\|P_{\Omega(x,\gamma)}(x - \nabla f(x)) - P_{\Omega(x,\gamma)}(x + \mu[h'(x)^T h(x) + \sum_{g_{i(x)>0}} g_i(x)\nabla g_i(x)])\| \leq \epsilon. \tag{10}$$

It remains to prove that

$P_{\Omega(x,\gamma)}(x + \mu[h'(x)^T h(x) + \sum_{g_{i(x)>0}} g_i(x)\nabla g_i(x)]) = x.$

To see that this is true, consider the convex quadratic subproblem

min y $\quad \|y - (x + \mu[h'(x)^T h(x) + \sum_{g_{i(x)>0}} g_i(x)\nabla g_i(x)])\|^2,$

s.t. $\quad y \in \Omega(x, \gamma),$

and observe that $y = x$ satisfies the sufficient KKT optimality conditions with the multipliers

$\lambda = -\mu h(x),$

$v_i = \mu g_i(x),$ for $g_i(x) \geq 0,$

$v_i = 0,$ else.

So, by (10),

$\|P_{\Omega(x,\gamma)}(x - \nabla f(x)) - x\| \leq \epsilon,$

as we wanted to prove.

Corollary 2.1

If x^* is a local minimizer of $(1)-(2)$ and if $\gamma \in [-\infty, 0]$, there exists a sequence $\{x^k\} \subset \mathbb{R}^n$ such that lim $x^k = x^*$ and lim $d(x^k, \gamma) = 0$.

Property 2.1

Assume that $g(x^*) \leq 0, h(x^*) = 0, x^k \to x^*$, and that, for some $\gamma \in [-\infty, 0], d(x^k, \gamma) \to 0$. Then, $d(x^k, \gamma') \to 0$ for all $\gamma' < 0$.

Proof

Consider the problems

$$\min \|x^k - \nabla f(x^k) - y\|^2, \tag{11}$$

s.t. $\quad g_i'(x^k)(y - x^k) \le 0, \qquad \text{if } g_i(x^k) \ge 0,$ (12)

$g_i'(x^k)(y - x^k) \le 0, \qquad \text{if } 0 > g_i(x^k) \ge \gamma,$ (13)

$h'(x^k)(y - x^k) = 0$ (14)

and

$\min \|x^k - \nabla f(x^k) - y\|^2,$ (15)

s.t. $\quad g_i'(x^k)(y - x^k) \le 0, \qquad \text{if } g_i(x^k) \ge 0,$ (16)

$g_i'(x^k)(y - x^k) \le 0, \qquad \text{if } 0 > g_i(x^k) \ge \gamma',$ (17)

$h'(x^k)(y - x^k) = 0$ (18)

Let y^k be the solution of (11)−(14). By the hypothesis, we know that $\|y^k - x^k\| \to 0$. Therefore, $y^k - x^*$. Let us know that, for k large enough, y^k satisfies the constraints (16)−(18). Let us prove that the same is true for (17). We consider two cases: $g_i(x^*) = 0$ and $g_i(x^*) < 0$.

If $g_i(x^*) = 0$, then for k large enough, $g_i(x^k) > \gamma$. Therefore, if in addition $g_i(x^k) < 0$, the constraint (17) is satisfied.

If $g_i(x^*) < 0$, then since $\|y^k - x^k\| \to 0$, we have that, for k large enough,

$g_i(x^k) + g_i'(x^k)(y^k - x^k) < 0.$ (19)

Therefore, the constraints (17) are also satisfied at y^k.

But y^k is the KKT point of the problem (11)−(14). By the analysis performed above, an active constraint of type (17) can correspond only to $g_i(x^*) = 0$ and $g_i(x^k) < 0$. In this case, $g_i(x^k) > \gamma$ for k large enough; thus the constraint is present in the set (13). This completes the proof.

Mangasarian – Fromovitz (MF) Constraint Qualification

If $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$, $\mu \ge 0$ are such that

$h'(x^*)^T\lambda + g'(x^*)^T\mu = 0$ and $\mu^T g(x^*) = 0$,then $\lambda = 0$ and $\mu = 0$.

Theorem 2.2

Assume that x^* is a feasible point. Let $\gamma \in [-\infty, 0]$. Suppose that there exists a sequence $x^k \to x^*$ such that $d(x^k, \gamma) \to 0$. Then, x^* is a Fritz-John point of (1)−(2).

Proof

Define

$y^k = P_{\Omega(x^k,\gamma)}\left(x^k - \nabla f(x^k)\right).$

So, y^k solves

$\min \ (1/2)\|y - x^k\|^2 + \nabla f(x^k)^T(y - x^k),$

s.t. $y \in \Omega(x^k, \gamma)$.

Therefore, there exist $\lambda^k \in \mathbb{R}^m$, $\mu^k \in \mathbb{R}^p$, $\mu^k \geq 0$ such that

$$\nabla f(x^k) + (y^k - x^k) + h'(x^k)^T \lambda^k + g'(x^k)^T \mu^k = 0, \tag{20a}$$

$$\mu_i^k [g_i(x^k) + g_i'(x^k)(y^k - x^k)] = 0, \qquad \text{if } \gamma < g_i(x^k) < 0, \tag{20b}$$

$$\mu_i^k [g_i'(x^k)(y^k - x^k)] = 0, \qquad \text{if } g_i(x^k) \geq 0, \tag{20c}$$

$$\mu_i^k = 0, \qquad \text{else.} \tag{20d}$$

Moreover, if $g_i(x^*) < 0$, we have that $g_i(x^k) < 0$ for *k* large enough; and since $\|y^k - x^k\| \to 0$, we have also that

$$g_i(x^k) + g_i'(x^k)(y^k - x^k) < 0$$

in that case. Therefore, we can assume that

$$\mu_i^k = 0 \qquad \text{whenever } g_i(x^*) < 0 \tag{21}$$

To prove the Fritz-John condition is equivalent to proving that the MF condition implies the KKT condition. So, from now on, we are going to assume that x^* satisfies the Mangasarian-Fromovitz (MF) constraint qualification.

Suppose by contradiction that (λ^k, μ^k) is unbounded. Define, for each *k*,

$$M_k = \|(\lambda^k, \mu^k)\|_\infty = \max\left\{\|\lambda^k\|_\infty, \|\mu^k\|_\infty\right\}.$$

Then,

$$\limsup M_k = \infty.$$

Refining the sequence (λ^k, μ^k) and reindexing it, we may suppose that $M_k > 0$ for all *k* and

$$\lim_k M_k = +\infty.$$

Now, define

$$\hat{\lambda}^k = (1/M_k)\lambda^k, \qquad \hat{\mu}^k = (1/M_k)\mu^k.$$

Observe that, for all *k*

$$\|(\hat{\lambda}^k, \hat{\mu}^k)\|_\infty = 1.$$

Hence, the sequence $(\hat{\lambda}^k, \hat{\mu}^k)$ is bounded and has a cluster point $(\hat{\lambda}, \hat{\mu})$ satisfying

$$\hat{\mu} \geq 0, \qquad \|(\hat{\lambda}, \hat{\mu})\|_\infty = 1.$$

Dividing (20) by M_k, we obtain

$$(1/M_k)[\nabla f(x^k) + (y^k - x^k)] + h'(x^k)^T \hat{\lambda}^k + g'(x^k)^T \hat{\mu}^k = 0$$

Taking the limit along the appropriate subsequence, we conclude that

$$h'(x^*)^T \hat{\lambda} + g'(x^*)^T \hat{\mu} = 0.$$

Together with (21), this contradicts the constraint qualification MF.

Now, since in (20), λ^k and μ^k are bounded, extracting a convergent subsequence, we have that

$\lambda^k \to \lambda^*$ and $\mu^k \to \mu^* \geq 0$.

Remark 2.2

We showed above that the AGP condition together with the Mangasarian – Fromovitz constraint qualification implies the KKT conditions. This is not true if the constraint qualification does not take place. For example, consider the problem

min x,

s.t. $x^2 = 0$.

The obvious solution of this problem does not satisfy the KKT conditions. On the other hand, it satisfies the AGP condition. Observe, for example, that the sequence $\{1/k\}$ trivially satisfies the AGP property.

Consider the nonlinear program given by

min x,

s.t. $x^3 \leq 0$.

Clearly, x^* satisfies the Fritz-John condition; but for any sequence $\{x^k\}$ that converges to $x^*, d(x^k, \eta)$ does not tend to zero.

SUFFICIENCY IN THE CONVEX CASE

As it is well known, in convex problems, the KKT conditions imply optimality. Here we will show the same sufficiency property which is true for the AGP optimality condition.

Theorem 3.1

Suppose that, in the nonlinear program (1)−(2), f and g_i are convex,

i = 1,…., p, and h is an affine function. Let $\gamma \in [-\infty, 0]$. Suppose that $x^* \in \Omega$ and that $\{x^k\} \subset \mathbb{R}^n$ are such that $\lim x^k = x^*, h(x^k) = 0$ for all k = 0, 1, 2, …, and lim $d(x^k, \gamma) = 0$. Then x^* is a minimizer of (1)−(2).

Proof

Let us prove first that

$\Omega \subset \Omega(x^k, \gamma)$, for all $x^k \in \mathbb{R}^n$.

Assume that $z \in \Omega$. *If* $g_i(z) \leq 0$ and $g_i(x^k) < 0$, by the convexity of g_i we have that

$$0 \geq g_i(z) \geq g_i(x^k) + g_i'(x^k)(z - x^k). \tag{22}$$

Moreover, if $g_i(z) \leq 0$ and $g_i(x^k) \geq 0$,

$$\begin{aligned} 0 &\geq g_i(z) \\ &\geq g_i(x^k) + g_i'(x^k)(z - x^k) \\ &\geq g_i'(x^k)(z - x^k). \end{aligned} \tag{23}$$

Therefore by (22) and (23), $z \in \Omega$ implies that $z \in \Omega(x^k, \gamma)$; so, $\Omega \subset \Omega(x^k, \gamma)$. Note that (5) holds since $h(x^k) = h(z) = 0$ and h is affine .

Let us define now

$$y^k = P_{\Omega(x^k,\gamma)}(x^k - \nabla f(x^k)).$$

Let $z \in \Omega$ be arbitrary. Since $z \in \Omega(x^k, \gamma)$ and y^k minimizes $\| x^k - \nabla f(x^k) - y \|^2$ on this set, we have that

$$\nabla f(x^k)^T(z - x^k) + (1/2) \| z - x^k \|^2$$

$$\geq \nabla f(x^k)^T(y^k - x^k) + (1/2) \| y^k - x^k \|^2.$$

Taking limits on both side of the above inequality and using the fact that $\| y^k - x^k \| \to 0$, we obtain

$$\nabla f(x^*)^T(z - x^*) + (1/2) \| z - x^* \|^2 \geq 0.$$

Since Ω is convex, the above inequality holds replacing z with $x^* + t(z - x^*)$ for all

$t \in [0,1]$; so,

$$t\nabla f(x^*)^T(z - x^*) + t^2(1/2) \| z - x^* \|^2 \geq 0.$$

Thus,

$$\nabla f(x^*)^T(z - x^*) + (t/2) \| z - x^* \|^2 \geq 0,$$

for all $t \in [0,1]$. Taking limits for $t \to 0$, we obtain that

$$\nabla f(x^*)^T(z - x^*) \geq 0.$$

Since $z \in \Omega$ is arbitrary, convexity implies that x^* is minimizer of (1)−(2).

THE STRONG CONDITION IN AGP

In the following lemma, we will show the strong AGP condition.

Lemma 4.1

For all $\gamma \in [-\infty, 0]$, if $x^k \to x^*$ and $d_s(x^k, \gamma) \to 0$, then $d(x^k, \gamma) \to 0$.

Proof

Observe that

$$\Omega(x, \gamma) \subset \Omega_s(x, \gamma), \qquad \text{for all } x \in \mathbb{R}^n, \gamma \in [-\infty, 0].$$

Let us call

$$w^k = P_{\Omega_s(x^k,\gamma)}\left(x^k - \nabla f(x^k)\right),$$

$$y^k = P_{\Omega(x^k,\gamma)}\left(x^k - \nabla f(x^k)\right).$$

Since $x^k, y^k \in \Omega(x^k, \gamma) \subset \Omega_s(x^k, \gamma)$, we have, by the definition of projections for all $t \in [0,1]$,

$$\| x^k - \nabla f(x^k) - w^k \|^2$$

$$\leq \| x^k - \nabla f(x^k) - (x^k + t(y^k - x^k)) \|^2$$

$$\leq \| t(x^k - \nabla f(x^k) - y^k) - (1-t)\nabla f(x^k) \|^2 \leq \| \nabla f(x^k) \|^2.$$

So,

$$(1/2) \| w^k - x^k \|^2 + (w^k - x^k)^T \nabla f(x^k)$$

$$\leq (t^2/2) \| y^k - x^k \|^2 + t(y^k - x^k)^T \nabla f(x^k) \leq 0,$$

for all $t \in [0,1]$.

Taking limits in the above inequalities and using the fact that $\| w^k - x^k \| \to 0$, we obtain that, for all $t \in [0,1]$,

$$\lim_{k\to\infty} (t^2/2) \| y^k - x^k \|^2 + t(y^k - x^k)^T \nabla f(x^k) = 0.$$

Therefore, for all $t \in [0,1]$,

$$\lim_{k\to\infty} (t^2/2) \| y^k - x^k \|^2 + (y^k - x^k)^T \nabla f(x^k) = 0. \quad (24)$$

Since $\| y^k - x^k \|$ is bounded, this implies that

$$(y^k - x^k)^T \nabla f(x^k) = 0.$$

Therefore, by (24) with $t = 1$, we obtain that $\| y^k - x^k \| \to 0$, as we wanted to prove.

Corollary 4.1

If x^* is a local minimizer of (1)−(2) and $\gamma \in [-\infty, 0]$, there exists a sequence $\{x^k\} \subset \mathbb{R}^n$ such that $\lim x^k = x^*$ and $\lim d_s(x^k, \gamma) = 0$.

Moreover, since by Lemma 4.1 $d_s(x^k, \gamma) \to 0$ implies that $d(x^k, \gamma) \to 0$, the following corollary follows in a straightforward way from Theorem 2.2.

Corollary 4.2

Assume that x^* is a feasible point. Let $\gamma \in [-\infty, 0]$. Suppose that there exists a sequence $x^k \to x^*$ such that $d_s(x^k, \gamma) \to 0$. Then x^* is a Fritz−John point of (1)−(2).

We will show that the strong AGP property is a sufficient condition for minimizers of convex problems, without the requirement $h(x^k) = 0$. This property is stated in the following theorem.

Theorem 4.1

Assume that, in the nonlinear program (1)−(2), f and g_i are convex, $i = 1, \quad ,p$, and h is an affine function. Let $\gamma \in [-\infty, 0]$. Suppose that $x^* \in \Omega$ and $\{x^k\} \subset \mathbb{R}^n$ are such that $\lim x^k = x^*$ and $\lim d_s(x^k, \gamma) = 0$. Then x^* is a minimizer of (1)−(2).

CONCLUSIONS

The optimality condition is proved explicitly in inexact restoration algorithms **[13−14]** and it can be easily verified in augmented Lagrangian methods. In other cases **[17−18]**, it could demand some detailed research analysis. The

set of AGP points is a sharper approximation to the set of local minimizers than the set of Fritz−John points; perhaps , there can be identified other sharp approximations to the minimizers that can be linked to the convergence of good minimization algorithms. This will be the subject of future research. Further research is necessary in order to extend the new optimality condition to nonsmooth optimization , variational inequality problems, bilevel programming , and vector optimization **[20−25]**. As in the smooth case analyzed in this paper, the guiding line comes from having in mind what efficient algorithms do for these problems, what kind of convergence results have already been proved for them, and which good and obvious practical behavior has been lost in those results.

REFERENCES

1. MANGASARIAN, O. L., and FROMOVITZ, S., THE Fritz−John *Necessary Optimality Conditions in Presence of Equality Constraints*, Journal of Mathematical Analysis and Applications, Vol. 17 , pp. 37−47 , 1967.

2. CASTELLANI, M., *Separation of Sets and Optimality Conditions*, Rendiconti del Circolo Matematico di Palermo, Serie II , Vol. 48 , pp. 27−38, 1997.

3. DIEN, P. H., MASTROENI, G, PAPPALARDO, M., and QUANG, P. H., *Regularity Conditions for Constrained Extremum Problems via Image Space*: *The Nonlinear Case*, Journal of Optimization Theory and Applications, Vol. 80, pp. 19−37, 1994.

4. GIANNESSI, F ., *General Optimality Conditions via a Separation Scheme*, Algorithms for Continuous Optimization, Edited by E. Spedicato, Kluwer, Dordrecht, Netherlands, pp. 1−23, 1994.

5. GIANNESSI, F ., and RAPCSAK, T., *Images, Separation of Sets , and Extremum Problems*, Recent Trends in Optimization Theory and Applications, Edited by R. P. Agarwal, Series in Applicable Analysis , World Scientific, Singapore, pp. 79−106, 1996.

6. GIANNESSI, F., *On Connections among Separation , Penalization , and Regularization for Variational Inequalities with Point-to-Set Operators*, Rendiconti del Circolo Matematico di Palermo , Serie II, Vol. 48, pp. 137−145, 1997.

7. BOGGS, P. T., and TOLLE, J. W., *A Strategy for Global Convergence in a Sequential Quadratic Programming Algorithm*, SIAM Journal on Numerical Analysis, Vol. 26, pp. 600−623, 1989.

8. FACCHINEI, F., JIANG, H., and QI, L., *A Smoothing Method for Mathematical Programs with Equilibrium Constraints*, *Mathematical Programming*, Vol. 85, pp. 107−134, 1999.

9. FERNANDES, L., FRIEDLANDER, A., GUEDES, M., and JUDICE, J. J., *Solution of Generalized Linear Complementarity Problems Using Smooth Optimization and Its Application to Bilinear Programming and LCP*, Applied Mathematics and Optimization, Vol. 43, pp. 1−19, 2001.

10. VICENTE, L. N., and CALAMAI, P., *Bilevel and Multilevel Programming: A Bibliography Review*, Journal of Global Optimization, Vol. 5, pp. 291−306, 1994.

11. BERTSEKAS, D. P., *Nonlinear Programming* , 2nd Edition, Athena Scientific, Belmont, Massachusetts, 1999.

12. FIACCO, A V., and McCORMICK, G. P., *Nonlinear Programming*, John Wiley, New York, NY, 1968.

13. MARTINEZ, J. M., *Two-Phase Model Algorithm with Global Convergence for Nonlinear Programming* , Journal of optimization Theory and Applications Vol.96, pp. 397−436, 1998.

14. MARTINEZ, J. M., and PILOTTA, E. A., *Inexact Restoration Algorithms for Constrained Optimization*, Journal of Optimization Theory and Applications, Vol.104, pp. 135−163, 2000.

15. ROCKAFELLAR, R.T., *Lagrange Multipliers and Optimality*, SIAM Review, Vol. 35, pp. 183−238, 1993.

16. DENNIS, J.E., EL-ALEM, M. and MACIEL, M.C., *A Global Convergence Theory for General Trust –Region-Based Algorithms for Equality Constrained Optimization,* SIAM Journal, Vol. 5, pp. 177−207, 1995.

17. PSIAKI, M. L., and PARK, S., *Augmented Lagrangian Nonlinear Programming Algorithm that uses SQP and Trust Region Techniques,* Journal of Optimization Theory and Applications, Vol. 86, pp. 311−326, 1995.

18. SPELLUCCI, P., *An SQP Method for General Nonlinear Programs Using Only Equality Constrained Subproblems,* Mathematical Programming, Vol. 82, pp. 413−447, 1998.

19. ANDREANI, R., and MARTINEZ, J. M., *On the Solution of Mathematical Programming Problems with Equilibrium Constraints*, Mathematical Methods of Operations Research, Vol. 54, pp. 345−358, 2001.

20. BIGI, G., and PAPPALARDO, M., *Regularity Conditions in Vector Optimization*, Journal of Optimization Theory and Applications, Vol.102, pp. 83−96, 1999.

21. GIANNESSI, F., MASTROENI, G., and PELLEGRINI, L., *On the Theory of Vector Optimization and Variational Inequalities: Image Space Analysis and Separation,* Vector Variational Inequalities and Vector Equilibria, Mathematical Theories, Edited by F. Giannessi, Kluwer , Dordrecht, Netherlands, pp. 153−216, 2000.

22. GIANNESSI, F., and PELLEGRINI, L., *Image Space Analysis for Vector Optimization and Variational Inequalities; Scalarization,* Advance in Combinatorial and Global Optimization, Edited by A. Migdalas, P. Pardalos, and R. Burkard, World Scientific, River Edge, pp. 982−1121, 1999.

23. FAVATI, P., and PAPPALARDO, M., *On the Reciprocal Vector Optimization Problem,* Journal of Optimization Theory and Applications, Vol. 47, pp. 181−193, 1985.

24. PAPPALARDO, M., *On Stationarity in Vector Optimization Problem*s, Rendiconti del Circolo Matematico di Palermo, Serie II, Supplemento 48, pp. 195−201, 1997.

25. PAPPALARDO, M., *Vector Optimization*, Encyclopedia of Optimization, Edited by P. Pardalos and C. Floudas, Kluwer, Dordrecht, Netherlands (to appear).

26. Brunato M. and Mascia F., Reactive Search and Intelligent Optimization, Operations Research/Computer Science Interfaces Series, Vol. 45, Springer, November 2008.

27. Yeganova **L.** and Wilbur, W.J. Isotonic Regression under Lipschitz Constraint, Journal of Optimization theory and Applications, 2009.

28. Tucker; A. W. "Linear and nonlinear Programming", Operations Researcg,5, no.2,1957, 244-257.

29. Kuhn H. and Tucker, A.W. "Nonlinear Programming," in Proceeding of the Second Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman, ed., University of California Press, Berkeley, Calif., 1951, 481-492.